

OVERALL SCORE

46 /100

Needs Work

Executive summary

This site has a useful foundation, but important gaps still limit AI readability. Key strengths include plain-text page access, while AI guidance file and sitemap need attention. Recommended next step: publish an AI guidance file as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Recommended next step

- 1. Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.
- 2. Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.
- 3. Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.

Signal breakdown

Crawlability

Warn 10/20

Serve a non-empty HTML homepage with a canonical link tag that points to the preferred public URL.

Robots.txt

Warn 15/15

Publish /robots.txt with explicit crawler rules and Sitemap references so search and AI crawlers can discover site policy.

llms.txt

Fail 0/15

Publish /llms.txt as text or markdown with more than 200 characters, markdown headings, and at least one absolute URL.

Sitemap	Fail	0/10
Publish a valid XML sitemap at /sitemap.xml and reference it from robots.txt so crawlers and AI systems can discover important URLs.		
Markdown support	Pass	15/15
The homepage returns markdown content when requested with Accept: text/markdown, giving AI systems a cleaner text representation with less navigation chrome to parse.		
Semantic HTML	Warn	5.7/10
Add exactly one h1 element that describes the page topic. Add missing semantic elements: main, article, nav, footer. Add visible body copy after script and style removal so the page has at least 200 words.		
Structured data	Fail	0/10
Add JSON-LD structured data with Organization or WebSite schema so AI systems can identify the site owner or website entity.		
Content signals	Fail	0/5
Add the standard directive 'Content-Signal: ai-train=no, search=yes, ai-input=yes' to robots.txt, HTML metadata, or HTTP headers so AI systems can discover content usage preferences.		

Suggested fixes

Fix Structured data

HTML

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@graph": [
    {
      "@type": "Organization",
      "@id": "https://www.unifeyn.app/#organization",
      "name": "Unifeyn - AI study workspace grounded in YOUR materials",
      "description": "Turn your notes, PDFs, slides and videos into grounded chat, study guides, quizzes, flashcards, mind maps and audio summaries.",
      "url": "https://www.unifeyn.app/",
      "logo": "https://www.unifeyn.app/apple-touch-icon.png"
    },
    {
      "@type": "WebSite",
      "@id": "https://www.unifeyn.app/#website",
      "name": "Unifeyn - AI study workspace grounded in YOUR materials",
      "description": "Turn your notes, PDFs, slides and videos into grounded chat, study guides, quizzes, flashcards, mind maps and audio summaries.",
      "url": "https://www.unifeyn.app/",
      "publisher": {
        "@id": "https://www.unifeyn.app/#organization"
      },
      "inLanguage": "en"
    }
  ]
}
```

Continued in the full scan report...

Fix Content signals

HTML

```
<meta http-equiv="Content-Signal" content="ai-train=no, search=yes, ai-input=yes" />
<meta name="content-signal" content="ai-train=no, search=yes, ai-input=yes" />
```

Full report

<https://llmscan.dev/scan/WZITcsjv9AunE3tEFUHqV>